APPENDIX F-3

Challenges in Projecting Human Health Impacts from Exposures to Perchloroethlyene

White Paper by Dr. Lorenz Rhomberg,
Gradient Corporation, Cambridge, MA

Challenges in Projecting Human Health Impacts from Exposures to Perchloroethylene

Lorenz R. Rhomberg, Ph.D.

Principal Scientist
Gradient Corporation
Cambridge, MA

1 Introduction

When analysis of the toxic effects of chemicals is applied to the task of assessing benefits of regulations that limit exposure, the goal is to estimate the impact of changes in exposure regimes on changes in the burden of ill health in the exposed population. This differs from the aim of traditional regulatory risk assessment, which is to define exposure levels that can be deemed "safe," or at least that can be found to pose no more than "acceptable" risks. That is, the usual methods seek to define dose levels without pronounced impacts, not to estimate or to characterize the impacts that may occur.

Not surprisingly, traditional methods are ill suited to the estimation and description of toxic effects to be expected when chemical exposures approach and exceed levels that can be assuredly ruled safe. An often-mentioned issue is that traditional risk assessment methods are "conservative" in that they deal with uncertainties in the inferential process by making estimates or assumptions unlikely to underestimate risk, thereby tending to overestimate risk, at least on average. Such biases distort the assessment of benefits gained from avoided exposure.

Two further issues are perhaps as important, however, and may be more difficult to remedy. First, owing to their focus on defining doses without unacceptable effects, existing approaches often say little about what specific toxic phenomena are to be expected at exposures exceeding "safe" levels. Second, because they focus on individual risks to benchmark individuals with defined "high-end" exposures, existing approaches are not geared to estimating population risks among a large group of subjects with varying levels of exposure that may fluctuate in time or consist of occasional high-exposure episodes.

To undertake benefits assessment, some new approaches to analysis of toxic effects will have to be considered. This paper attempts to examine the challenges and to consider modifications to risk analysis methods that may help to address some of the questions. To ground the discussion in the context of data that are available for actual toxic agents, the example of perchloroethylene is used.

It is best to begin by defining goals, even if they represent ideals that we are unlikely to achieve in practice. In order for economic analysis to measure the benefits of regulations that restrict exposure to a potentially toxic agent, it must have estimates of the burden of ill health in the exposed population as it would be expected to exist when the regulation is applied as well as when it is not applied. The differences between these constitute the avoided health impacts, and the values placed on them (which, thankfully, it is not my task to address) largely constitute the benefits of the regulation. Clearly, at least one of these scenarios (with the regulation or without) is hypothetical, and so even in the ideal case we cannot rely solely on observation. Modeled projection of health impacts to be expected in a population under hypothetical exposure scenarios is a necessary part of the analysis.

1.1 Needs

It would seem that the ideal toxicological analysis would provide characterization of the following:

1. *What* specific responses are engendered by exposures to the toxic agent?
2. For responses that are graded, *how severe* is the response? How does severity progress over time?
3. *When* (in the course of an ongoing exposure) do responses arise? *How long* do ill effects endure?
4. *In whom* do responses arise?

These bear some discussion. First, to put value on a case of toxicity avoided, it helps to specify the effect in question, since different effects (and different severities) have different impacts on the quality of life. Exiting methods typically eschew making statements about the specific nature of toxic effects in humans that are extrapolated from animal studies. Animal carcinogenicity is assumed to indicate a human risk for some type of cancer, but this is not necessarily expected to manifest itself as the same type of cancer seen in the animals. Noncancer toxicity assessments define doses that appear to avoid all adverse responses seen among experimental animals, and the most sensitive of these is deemed the "critical effect," but it is not specified which effects are to be expected in humans in exposures that exceed "safe" levels. Ideally, then, methods for benefits assessment should aim at making more specific projections about the nature of the toxicity to be expected in sufficiently exposed humans. They should recognize that several toxic effects may be at issue, not solely the one that was used to set the acceptable dose in the regulation being examined.

Similarly, existing methods do not project when during the course of an ongoing exposure the adverse effects will become manifest. To place a value on a case of toxicity, however, one would want to know when in life it appears, how long the state of ill health endures, whether it changes in severity, whether the disease fully or partly regresses upon cessation of exposure, how much the length of life is shortened,

and how a change in exposure at some midlife point (resulting, say, from the imposition of a regulation) changes the likelihood of response. In short, one would like not just dose-response relationships, but descriptions of response as a function of dose-rate and time, including description of the consequences of non-constant dose rates. Exposure-dependence of the course of disease in any cases engendered is also of interest.

It should be clear that the concern is for population risks, not just individual risks (which are often the focus of traditional assessments). We seek to characterize all the effects as they are (or would be) realized in an actual human population of interest. The hypothetical 70-year fence-line resident, the pica child, the worker laboring 45 years at a degreasing tank, and other standardized individual scenarios of exposure that define benchmarks of individual risk in regulation-setting assessments are not at issue, not just because they are "high-end" exposures but because they represent but a few individuals among the many in the population whose collective benefits we wish to address.

Ideally, we would want to describe not only the full frequency distribution of exposure levels, but also when and by whom the various exposures are experienced, since exposures at different ages or in different patterns over lifetime will differently affect the likelihood of responses (and we may wish to place different value on responses occurring at different times or in people with different prior states of health). Multiple sources and pathways of exposure exist, and people change their geographic locations and local exposures on timescales ranging from minutes to years. When a regulation is phased in, or when an agent persists in the environment even after controls are imposed, the exposures will change year by year, and this time pattern may be important to characterize to gauge accrual of benefits from the exposure restriction.

These facts make for major challenges to exposure assessment. (Exposure methods are not my focus, but the issues should not be overlooked.) Creating a complete inventory of the individual histories of exposure in an entire diverse population may seem a daunting task, but considerable progress has been made in approaching such a description using simulation modeling. In this approach, the events and settings that lead to exposure in a population are described as random variables, and a large set of simulated life histories can be assembled (ILSI 1998) that describes the diversity of experiences in a whole population.

1.2 Uncertainty

Existing methods in risk assessment for projecting human risks from experimental observations of toxicity in animals are highly uncertain. Even use of epidemiological studies entails uncertainty in characterization of exposures, in description of responses, and in generalization from the study population to the more general population of interest. New methods that attempt to make more detailed pronouncements regarding the nature of endpoints and the timecourse of their manifestation while acknowledging the complexity of the distribution of human exposures are bound to be still more uncertain.

Any demand that an analysis of benefits cannot be undertaken until impacts of exposure can be projected with confidence dooms the enterprise. It also misses the point. While we do our best to project outcomes with precision, uncertainty cannot be avoided, only characterized. It is the fact that outcomes are uncertain that makes them risks. The assessment of the costs and benefits of regulation can be regarded as a problem in decision under uncertainty—we have to decide how much to spend to control exposures in the face of uncertainty about how much benefit (in terms of reduced health impact) we will receive. The decision to incur regulatory costs is deemed a good one for society if the mathematical expectation of

the uncertain benefits exceeds that of costs. The expectation is not the single most likely value, but rather the average over possibilities, each weighted by its likelihood of being true.

Seen from this point of view, the characterization of uncertainty in the projection of health effects is central to the analysis. The characterization of risk consists of specifying an array of possible outcomes or courses of events, each element of which is associated with the likelihood of its occurrence and the consequences should it indeed occur (Kaplan and Garrick 1981). In the present case, the likelihoods constitute our relative confidence in the alternative projections of health impacts. We want to avoid the "upper bound" and "worst-case" nature of much of existing methodology, but at the same time we should not seek only single "best estimates" (such as maximum likelihood curve fits). Instead, we should seek to characterize the distribution of possibilities.

2 Perchloroethylene

Perchloroethylene ("Perc," CAS No.127-18-4) is a high production-volume chlorinated solvent used as a chemical intermediate, as a solvent and degreasing agent, and as the primary solvent in drycleaning operations. Perchloroethylene is moderately volatile; without containment and measures for vapor recovery, use and disposal can result in considerable release of vapor to the atmosphere. Spills and leaks during storage have resulted in cases of contaminated soil and groundwater. Because the resulting exposures to workers and the general public lead to concerns for potential human health effects, the use and disposal of perchloroethylene is subject to regulation aimed at limiting workplace concentrations and releases of the chemical to the environment. The mandated controls can be costly, and it is of interest to establish how much impact on the health of the human population is avoided through their application.

A full review of the exposures to perchloroethylene and a complete characterization of its toxicologic and epidemiologic database are beyond the scope of this paper. The following overview, drawn from IARC (1995), EPA (1985, 1991), ATSDR (1997), OEHHA (1999) and other sources, gives a perspective on the available information that is sufficient for the present discussion of methodological issues.

Exposure: Worldwide annual production of perchloroethylene (which has declined somewhat in recent years) is in the hundreds of thousands of tons. Some 55% is used for drycleaning, 23% as a chemical intermediate (mostly for CFC production, which is declining), and 13% for liquid and vapor degreasing, with other uses including fabric treatment and paint stripping. Sampled air concentrations vary considerably in degreasing facilities, but means are often on the order of 10-100 ppm (parts per million) with some individual air samples in the 1,000 ppm range. Occupational exposures in drycleaning facilities are on the order of 10-50 ppm (IARC 1995).

Ambient air levels are much less and are reported here in parts per billion (1000 ppb = 1 ppm); they vary somewhat with season and are generally higher in urban than in rural air. Levels of 0.2 to 2 ppb are usually found in outdoor urban air. Indoor levels are often higher, sometimes tenfold outdoor levels. Peak levels in apartments above drycleaning establishments have been measured at 1000 ppb and higher. Off-gassing from drycleaned clothes can lead to temporarily high levels in automobiles (about 2000 ppb but with reports up to 300,000) and in homes (about 400 ppb).

Pharmacokinetics and Metabolism: Perchloroethylene is readily absorbed after inhalation or ingestion. Much of this is exhaled unchanged, but on the order of 30-50% is metabolized at low exposure levels in both rodents and humans. Most of this metabolism is *via* an oxidative pathway, but at exposure levels higher than about 100 ppm (such as in rodent lifetime bioassays), the oxidative pathway becomes increasingly saturated. This leads to proportionally higher metabolism by a glutathione-conjugation

pathway (which is still small in absolute terms) and higher exhalation of unmetabolized compound. Glutathione conjugates can be further metabolized in kidney to reactive, apparently genotoxic compounds, but oxidative metabolites (and perc itself) do not appear to be genotoxic. Several pharmacokinetic models of perchloroethylene metabolism exist; they agree in broad outline but differ in detail, especially regarding rodent-human differences in the extent of the conjugative pathway.

2.1 Observations of Toxicity in Humans

Neurological effects have been seen in populations occupationally exposed to bouts of high perchloroethylene concentration in air, and the nervous system seems to constitute the most susceptible target in humans. Overt symptoms such as headache, nausea, and ataxia are not seen in experiments at doses below 100 ppm, and these are fully reversible. Subtler pre-clinical neurophysiological and neurobehavioral effects such as changes in electroencephalograms, visual-evoked potentials, color vision discrimination, and tests of coordination or reaction time show detectable influence of exposure at levels between 15 and 100 ppm, although these, too, are reversible upon cessation of exposure. No clear evidence suggests permanent neurological effects from chronic occupational exposure, but some studies report detection of significant differences in memory and reaction time.

Case studies exist of workers exposed to very high levels in industrial accidents (*e.g.*, a worker found unconscious in a pool of solvent) in which serious liver or kidney damage occurred, but in such cases there is apparent full recovery within weeks. As with neurological effects, subtler pre-clinical changes that are considered markers of potential toxicity are seen in some studies of workers with exposures in the 20-30 ppm range. Many of these are elevations in serum concentrations of certain liver-cell enzymes (SGOT, SGPT, GGT) that are taken to signal some loss of integrity or increased permeability of liver cells, and hence possible beginnings of hepatotoxicity. It is typical for these quantitative measures to be within normal range in all subjects yet the means for exposed and unexposed groups are statistically different.

Some studies suggest slightly increased rates of spontaneous abortion or menstrual complaints in women with occupational exposures, and some studies suggest longer times to conception in couples with one or the other parent exposed. No associations with stillbirth, low birthweight, or malformations have been seen.

Several occupational epidemiological studies of carcinogenic effects have been conducted of drycleaning workers and those exposed in settings where degreasing activities lead to elevated air concentrations. Various inconsistent small elevations of one or another type of tumor have been reported (lymphopoietic, female genital, bladder, kidney, breast), but the only one showing any consistency is esophageal cancer. This effect (SMR 2.1 and 2.6) was seen in two drycleaning employee cohorts (but only in black men in one of them). A case-control study of esophageal cancer showed a non-significant association with employment in drycleaning. Esophageal cancer is subject to influence by smoking and alcohol use. Moreover, perchloroethylene is not the only chemical exposure for many of the workers in these studies.

2.2 Observations of Toxicity in Experimental Animals

Many of the noncancer effects seen in humans are seen in animals as well, but often at higher doses and for more overt and frankly toxic versions of the effect (since subtle effects are difficult to detect). Thus, animals acutely exposed to over 1000 ppm showed ataxia and anesthesia as well as altered psychomotor functions. Effects on brain weight were seen in rats at 600 ppm for 4 or 12 weeks. High exposures also produce liver and kidney toxicity, and the biochemical markers such as serum enzymes also appear at

exposures on the order of 25 ppm. Some effect on litter size and survival during lactation were seen at 1000 ppm.

In lifetime carcinogenicity bioassays, perchloroethylene by gavage (NCI 1977) and by inhalation (NTP 1986) increased hepatocellular carcinomas in male and female mice. The NCI study has been questioned because the perchloroethylene used was stabilized with epichlorohydrin, itself an animal carcinogen. Inhalation in rats led to an increase in mononuclear cell leukemia in both sexes, although response was no higher at the high than at the low dose. In addition, treated male rats had a few renal tubular cell adenocarcinomas that, although not statistically elevated compared to controls, were considered toxicologically significant owing to their historical rarity.

None of these animal tumor responses is without some controversy regarding its applicability as indicator of potential human risk. Mice of the strain tested are particularly prone to such tumors, and they appear at high levels even in controls. A major metabolite of perchloroethylene, trichloroacetic acid, induces proliferation of peroxisomes in mouse liver cells at high doses, and the damage or oxidative stress these cause may be involved in the induction of tumors, although other evidence questions the role of peroxisomes in hepatocarcinogenesis and the correlation of their induction with liver tumor induction has counterexamples. Humans have very little peroxisomes induction, even at high exposures, and the background rate of liver cancer is much lower than seen in mice. Meanwhile, trichloroacetic acid administered to mice in drinking water or experienced as a metabolite of trichloroethylene (which is similar in toxicology and metabolism to perchloroethylene) causes similar liver tumors at doses below those inducing peroxisomes and without inducing evident cell proliferation.

Similarly, the rat strain studied is prone to mononuclear cell leukemias, a tumor type with no clear analogue in humans (it is splenic, and human leukemias originate in marrow). The rat controls have high responses, although the rate is observed to vary among studies. Male rats can develop kidney tumors from some chemicals that inhibit degradation of a male rat-specific protein (α_{2u}-microglobulin) that accumulates in renal tubule cells, causing toxicity. This syndrome is unique to male rats and is considered irrelevant to human risk (since humans lack the mechanism altogether). Perchloroethylene metabolites appear to cause this phenomenon in male rats, but only at doses higher than those in the NTP bioassay, suggesting that a different mechanism is responsible. On the other hand, bioassay-level exposures to perchloroethylene do induce kidney toxicity, probably as a result of the kidney's ability to further metabolize products of the conjugative pathway into reactive compounds (which also may be genotoxic). But this phenomenon, including the kidney toxicity, is seen in mice as well, and mice do not have elevations in kidney tumor risk. There is evidence that the conjugative pathway and the activation of metabolites in kidney happen in humans, but the quantitative extent is unclear.

3 Projecting Cancer Risks

If we want to assess the benefits of limiting perchloroethylene exposure in terms of avoiding cancer risks, the first question to face is whether perc is a human carcinogen at all. One possible stance is to conclude that evidence is insufficient to treat this compound as a human carcinogen, and hence there is no cancer risk among exposed people (and thus no benefit from restricted exposure). Even if we feel that this is the single best-supported conclusion, however, there is some probability that we are wrong, and if we are, the cancer risk that may exist is overlooked. By the same token, it would be a mistake to put all our credence in an analysis that assumes that perc is a human carcinogen, ignoring the substantial probability that any risks so calculated are illusory.

Current methods force just such an either-or decision, with the decision process couched in the weight-of-evidence determination in hazard identification. In the case of perchloroethylene, the weight-of-evidence regarding human carcinogenicity is particularly muddled. IARC (1995) has called perc a 2B "probable human carcinogen" based on what it judges to be "limited" epidemiological evidence and "sufficient" animal evidence. EPA has withdrawn a former B2 classification, and the SAB has declared perc to be on the borderline between B2 and C. For the purposes of benefits assessment, our purpose should not be to resolve the hazard question, but to figure how best to hedge our estimates of cancer risks to account for the ambiguity. At present, there is no rigorous analytical scheme for doing this, so we need to rely on some kind of expert judgment. For sake of argument, I propose to put 10% weight on the possibility that perchloroethylene does indeed pose a human cancer risk (at some levels of exposure relevant to the assessment), and 90% on the possibility that it does not. My judgment attempts to account for the inconsistency of results among epidemiological studies, the likelihood that exposures to other agents or confounding by smoking or alcohol apply, the inconsistency among animal cancer results and lack of concordance with observations in humans, and the lack of biological hypotheses for why esophageal cancer in particular should be caused by perchloroethylene.

A variant of this approach would be to make separate judgments about each potential basis for a human cancer risk estimate, *i.e.*, a judgment about the esophageal cancer, about the bladder cancer, about the hematopoietic cancers, *etc.* seen among the human studies, as well as judgments about the liver cancer, leukemias, and kidney tumors in the animal studies. Each weight could then be multiplied by the study-specific estimate of risk (made contingent on its presumed relevance). This appropriately allows some (very small) probability that, say, both the mouse liver tumors and the human-study bladder tumors are indicating some actual human cancer risk from perc.

At this point it is probably wise to emphasize the distinction between using such a hedging approach for setting a regulation in the first place and for estimating the benefits of a regulation set by some other reasoning. In my view, some degree of conservatism and precaution in setting allowable exposures is legitimate. What we get for our money is not just the reduction of health impacts, but some degree of assurance that we have done enough to protect public health. Nonetheless, when the question is the estimation of what the regulation has accomplished, what is needed is our best attempt to make objective estimates of the relative likelihoods that various levels of benefit have been achieved. Such an analysis informs not only the expected benefits (the mean over possibilities) but also the assessment of how much assurance we have in fact achieved.

The next question is to ask what the cancer potency is in humans, contingent on our provisional consideration that there is one. The problem most often pointed to in this realm is that current methods for describing dose-response relationships define "upper bound" risks rather than central estimates. As noted previously, the solution is not to use the single best-fitting dose-response equation (the maximum likelihood estimate), since this fails to express the variety of more-or-less reasonable dose-response relations and does not in general reflect the expected value of the risk.

Instead, a useful approach is to conduct a bootstrap analysis of the dataset. In this simulation-based approach, a large number of alternative datasets are generated by resampling the original data (with replacement), and a best-fitting curve is generated for each iteration. This expresses the variation in low-dose potency to be expected as a result of the experimental error inherent in a limited number of observations, and the mean of the distribution gives an unbiased estimate of the expected value over the various possible values, with each possibility weighted by its likelihood of occurrence.

Although this kind of experimental error is what is allowed for in the upper-bound calculations of traditional methods, such error does not constitute the only, or even the primary, source of uncertainty in estimates of human low-dose cancer potency. There are many analytical choices made in the projection of animal-based risk estimates to humans. Notably, these include the choice of dose-response model to fit (which ideally should reflect understanding of underlying biological mechanisms of the agent's toxic action) and the means for determining the toxicologically equivalent exposures in the experimental animals and in humans. Such factors should be thought of as aspects of model uncertainty, since they reflect not alternative realizations of some underlying distribution, but rather our uncertainty as to what structure for the analytical approach gives the best projections.

Several studies have attempted to address this kind of uncertainty by an extension of the "hedging" approach described above. Each analytical choice is expressed as a stated set of alternatives, and the alternatives are then given weights to reflect the perceived relative plausibility of the approaches they embody. In a simulation approach, one can then iterate the analysis many times, each time choosing one of the alternatives for each factor with a likelihood proportional to the weights they have been assigned. The resulting distribution of outcomes gives a description of the array of possible overall analytical answers and their relative plausibility. McKone and Bogen (1992) applied this approach to perchloroethylene cancer risks from contaminated drinking water, although they gave equal weights to all the alternative datasets and analytical methods considered. Thompson and Evans (1997) built on this approach to consider cancer risks from perc use in drycleaning. A major advantage of such analysis is that it allows examination of the contribution to overall uncertainty from the various components, and it lends itself to value-of-information analysis that seeks to define how investment in research efforts to reduce key uncertainties can be expected to pay off in informing regulatory decisions. (They found that the expected value of perfect information about perchloroethylene's potency exceeds that about exposures.)

Evans *et al*. (1994) applied a more extensive version of this approach to the description of the carcinogenic potency of chloroform. They used a panel of experts to provide weights on the various analytical choices, and they allowed for the weights placed on alternatives for one factor to be contingent on choices for other factors. They found a wide but not unreasonable distribution of implied potencies. The then-existing EPA potency estimate fell at a high percentile of the estimates, as is appropriate for an upper bound, but the whole distribution provides perspective on the expected amount of benefit that limits on chloroform exposure could be thought to achieve.

This process of elaboration of possible alternatives could be drawn out indefinitely, so one has to devise an approach that captures the main sources of uncertainty and describes them adequately for the purposes at hand. In the case of perchloroethylene, we have several (poor) choices of datasets to analyze (and hence a large weight on the notion that none of them is applicable), several alternative pharmacokinetic models, each of which could be subject to a characterization of the uncertainty distribution of its estimates of values of several different dose measures (reflecting different perchloroethylene metabolites in different tissues, in mice, rats, and humans), with different dose-response approaches to be considered in view of judgments about mechanism of carcinogenic action. Clearly, the approach is not easy to implement, but simplified versions could be used to give a reasonable view of the uncertainty about projections of human cancer risk.

Once we have such projections, we need to deal with the fact that they are unspecific about the kind of cancer to be expected in humans as well as the time of appearance of any tumors that are in fact caused. At present, there is no very satisfactory method for specifying these, but it is worthwhile considering how important it really is to do so. If we assume that most cancers have roughly similar impact on length and

quality of life, and if we assume that induced cancers appear with the same distribution over ages as the general burden of background cancers, we will probably not be far off.

4 Projecting Noncancer Health Effects

Many of the issues just discussed regarding cancer risks apply to noncancer risks as well, but there are some additional questions to be considered.

First, "noncancer" toxicity is a catchall category, and a single chemical may cause several different kinds of noncancer effects. In the traditional assessment process, a critical effect is identified as the basis of setting an exposure below which no adverse responses are expected, but above such a level, various toxicities may be caused, and as doses increase, the number of endpoints that may become important may increase, as effects with higher and higher population thresholds come into play. For instance, moderately high doses of perchloroethylene may cause neurological effects, and still higher ones may cause these plus renal toxicity. We must therefore keep in mind that a series of parallel endpoint-specific assessments is necessary, and not just an assessment of the endpoint on which the RfD is based.

Second, noncancer endpoints vary considerably in their severity. This is always part of the debate about "adversity" that arises when one is defining the critical effect. A benefits assessment must consider the fact that avoidance of some effects that are not frankly adverse may nonetheless have some value (albeit less than might be ascribed to a more severe effect). It may be legitimate, therefore, to assess endpoints that would not be considered a basis for an RfD, but nonetheless affect quality of life. For example, avoidance of headaches and dizziness from perc inhalation may be validly considered as benefits of regulation of workplace levels, even if they are not strictly "toxic" effects.

Third, since severity can vary a good deal, it becomes especially important to identify the nature of the toxic effects that may be engendered. As with cancer assessment, traditional methods do not specify what effects may be risked at doses above those deemed "safe," and it is not generally presumed that humans will have the same toxic effects as those seen in experimental animals, but such presumptions are necessary for benefits analysis to gain specificity.

Fourth, unlike cancer, the severity (and not just the frequency) of response increases with increasing dose. Much toxicity data is expressed in quantal form (with or without an effect of a given grade), and the increasing health impact of higher doses on those individuals showing effects may not be readily described. Since people vary in their tolerance of exposures to agents, at some doses, some individuals will respond and others will not. At higher doses, more people in an exposed population will respond, but those who already responded at a lower concentration will have more severe effects at a higher one. As a consequence, the mix of severity of responses will vary with dose.

Fifth, unlike cancer, which once started becomes autonomous and independent of the dose that caused it, noncancer effects may (or may not) be dependent on continued exposure. For exposures that can be avoided, and for toxic effects that become evident with relatively short latency, it may be that sufferers of moderate symptoms remove themselves from exposure and limit the impact on their health. (Of course, the need to do so might be considered a non-health impact to which value might be ascribed.) For example, someone experiencing mild neurological symptoms from perc exposure on the job might seek reassignment. On the other hand, an effect on a pregnancy outcome provides no opportunity to detect and avoid a developing problem.

Sixth, in a similar vein, different endpoints will have different latencies, typical durations, tendencies to progress or resolve, and different degrees of recovery or reversibility being possible. The impact on quality of life will depend heavily on whether the effect appears early or late in life, whether it is permanent or reversible, and whether it gets worse with time, with or without continued exposure. These are not matters treated in traditional assessments of noncancer risk.

Seventh, many noncancer endpoints are defined and measured in terms of markers or indicators of effects, but the endpoints themselves are not the primary concern. For example, the effects of perchloroethylene on finger-tapping frequency or color discrimination are examined because these objective tests are thought to be measurable manifestations of underlying neurological impacts. The benefit of restricted exposure is not in better finger-tapping ability or fineness of color discrimination, but in freedom from the underlying neurotoxicity that these markers are presumed to reflect. The quantitative connection of marker effects with the impairments of the underlying system being affected are not always very clear.

Eighth, the reason that differing levels of response are seen at different doses for effects presumed to have a threshold is that different individuals have different tolerances, or individual thresholds, or degrees of reserve capacity. Those who respond at the lowest doses will be those in the population with the least reserve capacity. It may be that such people are very nonrandomly distributed over age and other demographic categories, and it may be that those prone to response are prone because of pre-existing ill health or marginal health, and their change in health state may be different than is assumed if effects are thought to fall randomly on the exposed members of a population.

Finally, traditional approaches to noncancer risk assessment make little attempt to characterize the quantitative changes in probability of response with changing dose levels. The focus is on finding NOAELs or benchmark doses—doses substantially without effect—rather than to map the shape of the dose-response relationship. Moreover, the means to extrapolate effects from animals to humans are not as well developed as for cancer assessment. The extrapolations are covered by "uncertainty factors" that act in part to make extrapolation corrections (to human equivalent doses or to particularly sensitive humans) and in part to allow for case-by-case uncertainty about how big an extrapolation correction to make. That is, the analysis is more of a safety assessment than a risk assessment, and impacts of exposures above the RfD are not readily characterized.

Methodological changes are needed that make noncancer risk analysis capable of explicit estimation and extrapolation. This requires separating the two roles of the uncertainty factors into unbiased estimates of extrapolation and additional allowances for uncertainty in those extrapolations. One promising approach is to use (in place of fixed uncertainty factors) empirical distributions over many agents of the magnitude of extrapolation needed. Baird *et al*. (1996) have explored such an approach. In current ongoing work, I and my colleagues (Sandra Baird, John Evans, Paige Williams, Andrew Wilson) are further developing this approach for the assessment of reproductive and developmental toxicity of ethylene oxide in humans. We use empirical distributions over many chemicals of species differences in toxicologically equivalent doses for noncancer effects as well as empirical information about interindividual variation in sensitivity to arrive at unbiased estimates (with characterization of uncertainty) of the human dose-response relationship. Such results are suitable for making estimates of impacts of exposures at different dose levels, including those above traditionally defined reference doses. The result of this analysis is a set of distributions of the uncertainty in doses expected to lead to different levels of response in an exposed human population, the kind of assessment that is needed for analysis of benefits of regulation for noncancer endpoints.

5 Conclusions

Risk analysis in support of benefits assessment is different in aims from analysis for the setting of regulatory levels as currently practiced. It needs to be focused on estimation of effects, not on the bounding of regions of exposure where one can be very confident that unacceptable impacts are not to be expected. Accordingly, methods of risk analysis for benefits assessment need to be somewhat different.

There are many profound challenges, but I have tried to show that they are approachable, at least in concept. The appropriate analyses are not quick or easy, and there is no minor tweak to existing methods that will make them fully applicable. Having laid out attempts to define the ideal analysis, perhaps simpler versions that are more readily conducted will become evident.

It is important to distinguish the task of estimating actual health effects (and the uncertainty about that estimation) from the task of setting regulatory levels. The difficulty in estimating benefits should be clear from the above discussion. A good deal of judgment is necessary, and there is likely to be controversy in specific cases about the weights to be put on alternative possible estimates of the health effects engendered by an exposure. This makes it difficult to use analysis of benefits and costs to define what acceptable exposure levels should be. This being said, there is value in using such analysis to gauge how much value is gained from regulation, and how much uncertainty there is about the magnitude of such gain.

Acknowledgements

Leslie Beyer and Eric Dubé of Gradient Corporation contributed to the summary of toxicity of perchloroethylene. Production of this paper was supported by the U.S. Environmental Protection Agency's Office of Air and Radiation under Order No. 0D-6263-NALX.

References

ATSDR (Agency for Toxic Substances and Disease Registry) 1997. Toxicology Profile for Tetrachloroethylene: Update. US Department of Health and Human Services, Public Health Service, Atlanta, GA.

Baird SJS, Cohen JT, Graham JD, Shlyakhter AI, Evans JS. 1996. Noncancer risk assessment: a probabalistic alternative to current practice. Human Ecol. Risk Assessment 2:79-102.

EPA (US Environmental Protection Agency) 1985. Health Assessment Document for Tetrachloroethylene (Perchloroethylene). (EPA/600/8-82/005FA) US Environmental Protection Agency, Office of Health and Environmental Assessment, Washington, DC.

EPA (US Environmental Protection Agency) 1991. Response to Issues and Data Submissions on the Carcinogenicity of Tetrachloroethylene (Perchloroethylene). Review Draft. (EPA/600/6-91/002A) US Environmental Protection Agency, Office of Research and Development, Washington, DC.

Evans JS, Gray GM, Sielken RI Jr., Smith AE, Valdez-Flores C, Graham JD. 1994. Use of probabalistic expert judgment in uncertainty analysis of carcinogenic potency. Regul. Toxicol. Pharmacol. 20:15-36.

IARC (International Agency for Research on Cancer) 1995. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol.63. Dry Cleaning, Some Chlorinated Solvents and Other Industrial Chemicals. World Health Organization, International Agency for Research on Cancer, Lyon, France.

ILSI (International Life Sciences Institute) 1998. Aggregate Exposure Assessment: An ILSI Risk Science Institute Workshop Report. Risk Science Institute, International Life Sciences Institute, Washington, DC.

Kaplan S, Garrick BJ. 1981. On the quantitative definition of risk. Risk Anal. 1:11-27.

McKone TE, Bogen KT. 1992. Uncertainties in health-risk assessment: an integrated case study based on tetrachloroethylene in California groundwater. Regul. Toxicol. Pharmacol. 15:86-103.

NCI (National Cancer Institute) 1977. Bioassay of Tetrachloroethylene for Possible Carcinogenicity. Publication No. 77-813. National Cancer Institute, Department of Health and Human Services. Bethesda, MD.

NTP (National Toxicology Program) 1986. Toxicology and carcinogenesis studies of tetrachloroethylene (perchloroethylene) (CAS No. 127-18-4) in F344/N rats and B6C3F_1 Mice (Inhalation Studies). National Toxicology Program Technical Report TRY 311. National Institutes of Health.

OEHHA (Office of Environmental Health Hazard Assessment) 1999. Public Health Goal for Tetrachloroethylene in Drinking Water. Review Draft. Pesticide and Environmental Toxicology Section, Office of Environmental Health Hazard Assessment, California Environmental Protection Agency.

Thompson KM, Evans JS. 1997. The value of improved national exposure information for perchloroethylene (perc): a case study for dry cleaners. Risk Anal. 17:253-271.